United States
Environmental Protection
Agency

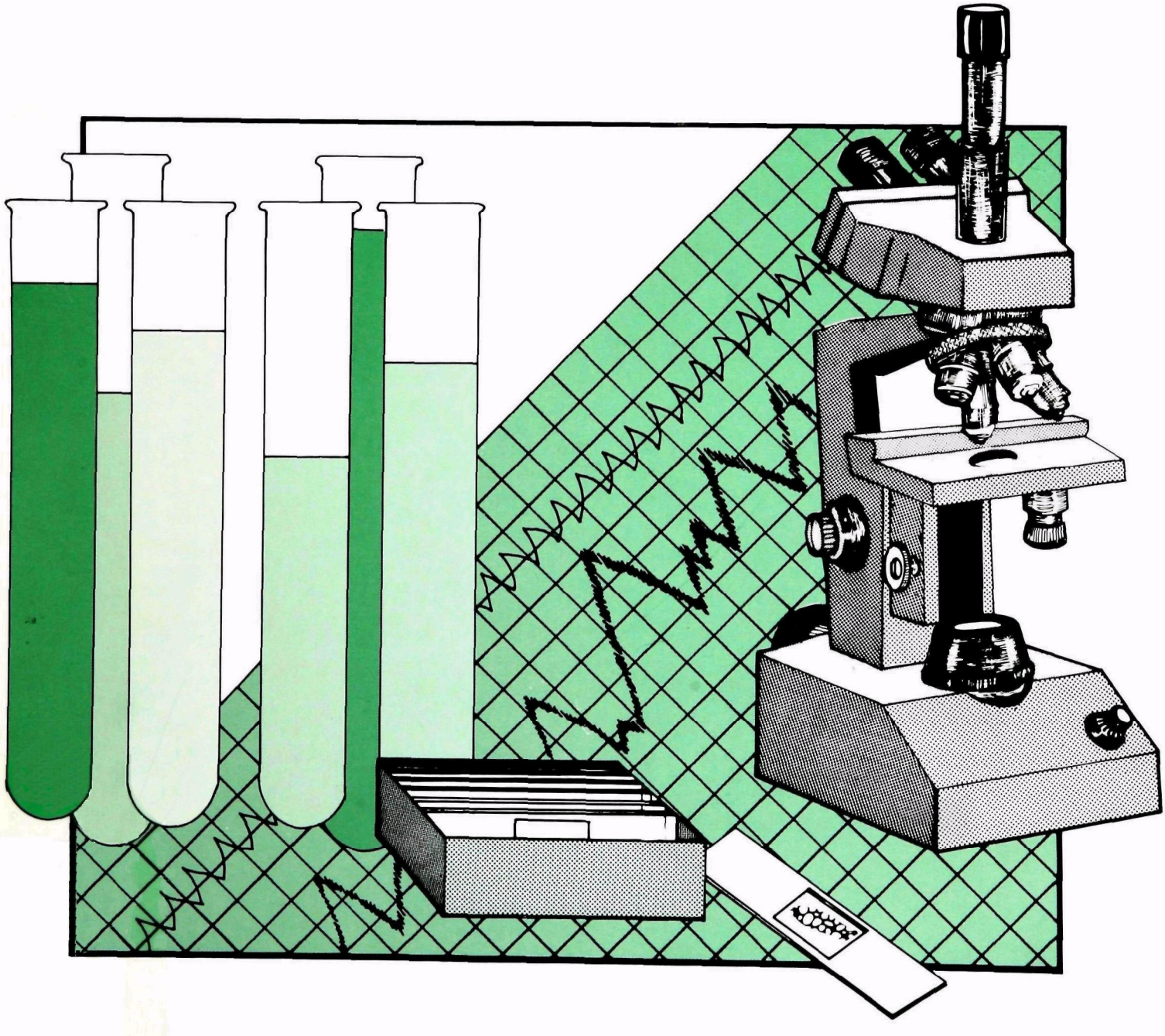
Office of Pesticide Programs
Washington, DC 20460

EPA-540/9-85-018
June 1985

EPA

Hazard Evaluation Division Standard Evaluation Procedure

Teratology Studies



EPA-540/9-85-018
June 1985

HAZARD EVALUATION DIVISION

STANDARD EVALUATION PROCEDURE

TERATOLOGY STUDIES

Prepared by

Laurence D. Chitlik, D.A.B.T.
Quang Q. Bui, Ph.D.
Gary J. Burin, M.P.H.
Stephen C. Dapson, Ph.D.

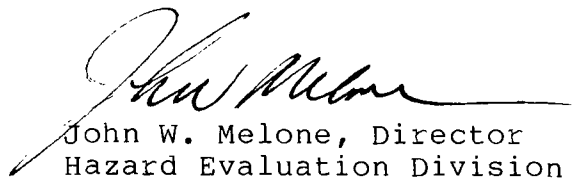
Standard Evaluation Procedures Project Manager
Stephen L. Johnson
Hazard Evaluation Division
Office of Pesticide Programs

United States Environmental Protection Agency
Office of Pesticide Programs
Washington, D.C. 20460

STANDARD EVALUATION PROCEDURE

PREAMBLE

This Standard Evaluation Procedure (SEP) is one of a set of guidance documents which explain the procedures used to evaluate environmental and human health effects data submitted to the Office of Pesticide Programs. The SEPs are designed to ensure comprehensive and consistent treatment of major scientific topics in these reviews and to provide interpretive policy guidance where appropriate. The Standard Evaluation Procedures will be used in conjunction with the appropriate Pesticide Assessment Guidelines and other Agency Guidelines. While the documents were developed to explain specifically the principles of scientific evaluation within the Office of Pesticide Programs, they may also be used by other offices in the Agency in the evaluation of studies and scientific data. The Standard Evaluation Procedures will also serve as valuable internal reference documents and will inform the public and regulated community of important considerations in the evaluation of test data for determining chemical hazards. I believe the SEPs will improve both the quality of science within EPA and, in conjunction with the Pesticide Assessment Guidelines, will lead to more effective use of both public and private resources.



John W. Melone, Director
Hazard Evaluation Division

TABLE OF CONTENTS

PREFACE

The purpose of this Standard Evaluation Procedure (SEP) for teratology studies is not to provide a complete discussion of the field of teratology but to act as a guide to the Regulatory Toxicologist for practical application of our current knowledge of this area. This SEP attempts to standardize, within the Office of Pesticide Programs, some of the common terminology employed in the evaluation of teratology studies and to provide useful information for risk assessment. For example, the term developmental toxicity is introduced in this SEP to more accurately describe the variety of end points that are of concern in a teratology study. It should also be noted that this document is consistent with the Agency's "Interim Guidelines for the Health Assessment of Suspect Developmental Toxicants."

The study of developmental toxicity of agents is rapidly developing and changing. Refinements in methodologies and techniques are almost a daily event and many issues and definitions basic to this area are still under debate. Further, no universal dictionary of terata is available and definitions of malformations and variations vary dramatically from laboratory to laboratory. Even new areas such as "behavioral teratology" are being rapidly re-defined as more inclusive fields such as "functional developmental toxicology."

The SEP points out some of the fundamental controversies in this area in order to caution the toxicologist. For example, the study reviewer should understand that findings in a study may either be transient or permanent. Where possible, methods of dealing with these types of issues are offered in the SEP. The toxicologist is also informed as to why the classical studies, as required by the Guidelines, sometimes fall short of resolving such issues and basic guidance is provided to the reviewer. Furthermore, it should be stressed that the toxicologist must maintain an understanding of the developmental toxicity issues faced by this program and, when necessary, follow up on these issues by requiring the most appropriate ancillary studies, e.g., post-natal studies.

Efforts have been made in this document to organize the subject into different sections: data acceptability, data evaluation and interpretation, and risk assessment.

The Data Acceptability section deals with specific questions regarding the adequacy of the experimental design, and the utility and the completeness of the data reported.

The two main end points of a classical teratology study, maternal toxicity and developmental toxicity (embryo/fetotoxicity) which includes teratogenicity, are discussed in the section on Data Evaluation and Interpretation along with the utility of developmental hazard indices. Factors influencing the outcome of toxicological manifestations such as diseases and methods of impregnation are addressed. Experimental animal data and attempts to differentiate malformations and variations are also discussed in this section. Effects of environmental agents on post-natal development are covered in the section on Functional Defects.

Another area discussed is risk assessment as it relates to developmental toxicants. The Office of Pesticide Programs has used the "Margin of Safety Approach" for a number of years. This approach considers the NOEL for developmental toxicity with the estimated exposure as a ratio. The only difference between this approach and the Safety Factor Approach is the a priori consideration of exposure data. Clearly, however, the area of developmental toxicity risk assessment needs considerable work and changes are to be anticipated in the near future.

This SEP should be considered as only a beginning and will require periodic updates.

TERATOLOGY STUDIES

I. INTRODUCTION

A. Definition

There are a variety of definitions for teratology or teratogenicity which may be found in the literature. One definition may be found under § 83-3 of the 1982 Subdivision F Guidelines: "Teratogenicity is the property of a chemical that causes permanent structural or functional abnormalities during the period of embryonic development." Teratology has also been defined as the study of abnormal development and congenital malformations. No matter what the definition, it should be understood that malformations are only one manifestation of developmental toxicity. A wide variety of end points are of interest in teratology studies. In addition to malformations and other structural alterations, embryo-fetal death, and other signs of developmental toxicity, are important indicators of an effect on the conceptus that should be assessed.

The period during which exposure takes place in teratology studies is usually the period of major organogenesis, but in some studies it may be extended from implantation of the conceptus up to parturition (Subdivision F Guidelines). The manifestations of developmental toxicity may be observed at any time during the life span of the test animal.

There is a lack of uniformity in the use of key terms in the field of developmental toxicology. For our purposes, the following definitions will be used:

Developmental Toxicology: The study of adverse effects on the developing organism which may result from exposure prior to conception (of either or both parents), during prenatal development (as in a classical Guidelines teratology study), or postnatally to the time of sexual maturation. Adverse developmental effects may be detected at any point in the life span of the organism. The major manifestations of developmental toxicity include: 1) death of the developing organism, 2) structural abnormality (malformations and variations), 3) altered growth, and 4) functional deficiency.

Embryotoxicity and fetotoxicity: The subset of developmental toxicity referring to adverse effects on the developing conceptus prior to parturition. The distinguishing feature between the two terms is the stage of development during which the injury occurred. Effects which occurred during the embryonic period are referred to as embryotoxicity and events during the fetal period are referred to as fetotoxicity. Oftentimes, embryotoxicity and fetotoxicity are difficult to distinguish and in those cases

the term embryo/fetotoxicity is appropriate. This subset of developmental toxicity includes: 1) in utero death, 2) structural abnormalities (malformations and variations), and 3) altered growth.

Teratogenicity: As defined in the 1982 Subdivision F Pesticide Assessment Guidelines § 83-3, "Teratogenicity is the property of a chemical that causes permanent structural or functional abnormalities during the period of embryonic development". In the Interim Guidelines for the Health Assessment of Suspect Developmental Toxicants, the term teratogenicity is only used to describe a "permanent structural change which may adversely affect survival, development or function." In reality, both definitions are consistent since all functional abnormalities must have a structural basis, even if only at the molecular level.

Altered Growth: Alterations (other than malformations) in the soft tissues, the skeletal system or body weight or size of the developing offspring. Altered growth may be induced at any stage of development and may be permanent or reversible.

Variations and Malformations: A variation is a divergence beyond the usual range of structural constitution but which may not adversely affect survival, development, or function. A malformation is a permanent structural change which may adversely affect survival, development, or function. Because development is considered as a continuum of events, there is no universally accepted classification for many structural alterations as either variations or malformations (See Section III.B of this document for examples of both malformations and variations).

Functional Developmental Toxicology: The study of the causes, mechanisms, and manifestations of alterations or delays in functional competence of the organism or organ system following exposure to an agent during critical periods of development pre- and/or postnatally.

Risk Assessment: The qualitative and quantitative determination of potential risk to humans.

The guidance provided in this document is intended to be used in conjunction with the Core Classification System for determining the acceptability of a teratology study for regulatory purposes.

In a regulatory setting, teratology studies are required for the purpose of attempting to determine the potential of an agent to induce developmental toxicity. The test data must be approached cautiously in attempting to extrapolate to humans due to fundamental differences in conditions of exposure and biological differences between the test species and humans. However, for obvious ethical reasons, regulatory agencies have little choice but to use test results from experimental animals treated under laboratory conditions to predict the risk of human developmental toxicity.

B. When Required

The Office of Pesticide Programs requires teratology testing in two species to support the registration of a product intended for food use (i.e., when tolerances or exemptions from tolerances are considered) and for nonfood uses if significant exposure of human females of child bearing age may reasonably be expected. Temporary tolerances generally require testing in only one species.

The reviewer must be aware that the above requirements are minimum data needs and that additional testing or modifications of routine protocols (such as the inclusion of a post-natal phase which may include functional assessment as well as information on the reversibility of altered growth) will sometimes be necessary for a more meaningful assessment of potential developmental toxicity. Structural similarity to developmental toxicants including known teratogens may trigger testing when it otherwise would not be required, but the reviewer should also be aware that small differences in molecular structure may lead to major differences in the potential for developmental toxicity and that structure-activity relationships in developmental toxicity testing are not well established. Nevertheless, the reviewer should consult published reference materials (e.g., Shepard, 1983; Schardein, 1985) and the Toxicology Branch files for structurally related compounds prior to determining data requirements for teratologic assessment if there is any doubt as to the need for testing.

The reviewer should also be aware that it is sometimes possible to waive the need for further teratogenicity testing if the "limit test" requirements in Subdivision F are met or exceeded. This section states that "if a dose level of 1000 mg/kg produces no evidence of embryo toxicity or teratogenicity, studies at other dose levels may not be considered. If a preliminary study at this high dose level, with definite evidence of maternal toxicity, shows no adverse effect on the embryos, studies at other dose levels may not be necessary."

C. Usefulness of the Teratology Study in a Regulatory Setting

Information gathered from teratology studies may be the only information that is available regarding the effects of *in utero* exposure or it may be supplemented by the results of a reproduction study. In either case, the teratology study yields important and often unique information regarding the effects of a compound on the conceptus. Furthermore, although it is the dam that is being dosed, the reviewer of a teratology study must focus on both the litter as well as the fetus. Maternal toxicity data is useful to confirm that appropriate dose-levels have been used to allow the greatest potential for eliciting an effect.

As noted previously, caution should be exercised in extrapolating animal developmental toxicology results to humans. However, these studies are the best indicators available in the prediction of potential adverse developmental effects in humans. Positive findings in one or more animal species triggers a regulatory response which consists of assessment of exposure (worker, dietary, or other) and a comparison (ratio) of these exposure levels with the No Observed Effect Level (NOEL) in the most sensitive species (Margin of Safety). The conduct of a developmental toxicity risk assessment is described in detail in Section IV of this document.

II. DATA ACCEPTABILITY

A. Range Finding Study

The purpose of the range finding study is solely to establish the proper selection of dose-levels for the primary teratology study. The information required in the range finding study should be adequate to assess maternal toxicity and should be consistent with that required in Section 80-4 of the Subdivision F Guidelines. Maternal weight gain during dosing and over the course of gestation is of great importance and data should be available for each animal on test. Gross necropsy results for each dam, detailed clinical observations, uterine and ovarian findings (such as the incidence of corpora lutea, implantations and resorptions), and food consumption data (if available) are necessary to fully assess maternal toxicity.

Developmental toxicity is not fully assessed in the range finding study and information relating to embryo/fetoxicity is usually limited to a rudimentary assessment of prenatal mortality and external fetal findings. Also, range finding studies often lack soft tissue and skeletal examinations of fetuses and utilize small numbers of litters and fetuses. Therefore, they are of limited utility for assessment of developmental toxicity, including teratogenicity.

The acute oral toxicity study in the same strain of animal may also provide useful information for the setting of dose levels. However, it should be recognized that the acute studies are the result of only a single dosing and that the sensitivity of a dam during pregnancy may be different than that of a nonpregnant animal. Other studies such as subchronic and metabolism studies may be useful in judging the appropriateness of dose-levels but rarely are adequate for this purpose.

In reviewing the range finding teratology study, the reviewer should pay particular attention to potential manifestations of maternal toxicity. Maternal deaths and increased incidence of abortions will often be observed at the higher dose-levels.

Although these are forms of maternal toxicity, dose-levels associated with a large number of abortions or maternal deaths are usually not optimal for the high dose in the primary teratology study (see Section IIIB). However, in some cases the margin between the dose-level inducing lethality or abortion and that associated with other forms of maternal toxicity such as decreased body weight gain (or absolute body weight loss), tremors or other manifestations of compound toxicity, may be small. In those cases, the deaths or abortions may be seen at the lowest levels which also induce other forms of maternal toxicity. Optimally, the spacing of the dose-levels should be adequate to select a high dose-level for the primary study which does not result in more than 10% maternal mortality.

A range finding study cannot be classified as more than Core Supplementary Data since it is not adequate for the assessment of developmental toxicity.

B. Primary Study

Information required to be in the primary study is clearly stated in "Subdivision F: Hazard Identification: Humans and Domestic Animals" Sections 80-4 and 83-3.

It is particularly important to be able to associate all maternal and fetal findings with individual animals. For example, it is necessary to be able to relate soft tissue, skeletal, and external findings to any individual fetus and to determine the dam from which the fetus was removed as well as the maternal toxicity data specific for that dam.

A frequent shortcoming in the reporting of a teratology study is the lack of detail regarding the visceral and skeletal examination techniques and findings. If there is any question regarding the adequacy of these examinations, the registrant should be requested to furnish detailed protocols. In many cases the laboratory which conducted the study can furnish a Standard Operating Procedure which describes in detail the method by which technicians are to conduct such examinations. It is also often necessary to request the registrant to furnish photographs of malformations, detailed definitions, and descriptions of findings and the grading of severity of findings. The need for this information as well as historical control data must be considered on a case-by-case basis.

Historical control data must be from the lab which has conducted the test and should be reported on both a study-by-study basis as well as collectively, with findings reported in terms of the number of litters with a finding per number of litters examined and the number of fetuses with a finding per number of fetuses examined. The historical data base should preferably span two years prior to and (when available) after the submitted study.

All historical studies should be dated, and if a vehicle control was used, that vehicle should be identified. Historical maternal body weight gain data is also useful.

III. EVALUATION AND INTERPRETATION

A. Evaluation of Study Conduct

The Subdivision F Guidelines identify reporting requirements for teratology tests. It is important for the reviewer to pay particular attention to the following:

1. Choice of Test Compound

The test compound should be that intended for commercial use which in most cases is the technical material. Testing of formulations is usually not required. Both the purity of the test compound and the identification of impurities are important information and must be requested if these data cannot be ascertained from the study report or the compound registration. The reviewer should be aware that exposure to impurities in the test material may be an important aspect of the potential for developmental toxicity of a compound (e.g., dioxins).

2. Study Protocol

The study protocol should be compared to that presented in Section 83-3 of the Subdivision F Guidelines as well as to the Core Minimum standards. Deviations from basic protocols are sometimes necessary due to the need for additional information. For example, a post-natal phase may be necessary to distinguish the reversibility of dilated renal pelvis (apparent hydronephrosis; Woo and Hoar, 1972) and true hydronephrosis. Histopathology or dual skeletal staining may sometimes be necessary to follow up on unusual findings. The method of visceral and skeletal examination is left up to the discretion of the testing laboratory provided that a thorough examination and a sufficient number of animals are used. For rats, mice, and hamsters, approximately equal numbers should be examined for both skeletal and visceral findings (Wilson, 1965). It is possible to examine all fetuses for both visceral and skeletal findings, but this is not required for rats, mice, and hamsters. Rabbit fetuses, being larger, are more easily examined both viscerally and skeletally and both examinations are in fact required for each rabbit fetus (Staples and Wilson techniques or modified techniques). In any case, the methods of visceral and skeletal examination should be clearly stated in the study report.

3. Choice of Vehicle

The vehicle should be appropriate for the delivery of the test compound, should not interfere with absorption, and should not

induce maternal or developmental toxicity. If there is question as to the toxicity of the vehicle, the registrant should be required to justify the choice of vehicle. A sham-control may sometimes be required if the potential toxicity of the vehicle is not completely understood.

4. Choice of Species

Generally, rats, mice, hamsters, or rabbits are used for testing. Studies conducted using other species may be acceptable but justification for the deviation should be requested from the registrant.

5. Route of Exposure

Dosing by gavage is the preferred route of administration in teratology studies because this route is considered to be closely related to human dietary exposure. However, studies conducted by other routes of administration, such as dermal and inhalation, are also acceptable to the Office of Pesticide Programs on a case-by-case basis if these routes are the primary situations for human exposure. Dietary administration is not preferred but may be considered acceptable if maternal toxicity is adequately demonstrated.

6. Reporting of Maternal and Fetal Findings

Reporting requirements have been stated in the Subdivision F Guidelines and will not be discussed in detail here. Clearly, however, it must be possible to associate all reported maternal and fetal findings with individual animals. It also must be possible to identify which dam showed any given clinical signs on any given day and to identify fetuses with each variation or malformation.

All reported mean data should be carefully compared to submitted individual data for possible inconsistency. The statistical methods used should be referenced and/or described. Statistical evaluations of all data including skeletal variation data should be available in the study report. The appropriate application of statistical methods should be verified.

7. Concurrence of Dosing of Test Groups

Dose groups should be run concurrently. The staggering of induction of pregnancy within dose groups is acceptable but the mean time of induction of pregnancy must not differ significantly from one dose group to another.

8. Final Reports

Study reports should be signed and dated and submitted as a final report. This, of course, is not necessary for reports published in the open literature. If a study report is not signed and dated, it must be assumed that it is subject to change and does not represent the final position of the investigator. It should be noted in the review of such a study that it is a draft and these studies should not be accepted as fully meeting regulatory requirements and classified as Core-Supplementary Data.

B. Study Interpretation

The following are the two major areas of a teratogenicity study that require careful assessment:

Maternal toxicity
Developmental toxicity

MANIFESTATIONS OF MATERNAL TOXICITY

The Subdivision F Guidelines specify that maternal toxicity must be observed in teratology studies. The actual wording reads:

> Unless limited by the physical/chemical nature or biological properties of the substance, the highest dosage level should include some overt maternal toxicity such as slight weight loss, but not more than 10% maternal deaths.

Compared to subchronic or chronic toxicity studies, teratology studies determine maternal toxicity based on a very limited number of parameters. These routinely include body weight, food consumption, clinical signs of toxicity, necropsy data, and reproductive data. Interpretation of maternal toxicity is further hindered by the short duration of exposure (during the period of major organogenesis) and the relatively small number of animals on test at each dose-level (especially in the case of rabbits).

1. Suitability of Test Animals

Among the mammals, rodents (rat and mouse) and rabbits have been used most frequently as animal models (Wilson and Fraser, 1978). Animals selected should be disease-free and of the same age and parity in order to eliminate variability in the results.

Basic husbandry information for commonly used animals is given as follows:

	Rat	Mouse	Rabbit
Gestation (days)	21+/- 1	20+/- 1	31+/- 1
Mating system	pair or colony†	pair or colony†	pair or artificial*

Source: Hafez, E.S.E., Reproduction and Breeding Techniques for Laboratory Animals eds, Lea & Febiger, Philadelphia, 1970.

† 1 male to several females
* Discussed in "Reproductive Status Section"

Pasteurellosis and coccidiosis are common diseases encountered in rabbits which may be manifested as: nasal discharge, diarrhea, and congested lungs, as well as pitted kidneys and brain lesions observed at necropsy (Benirschke et al., 1978). A high incidence of such observations is suggestive of a questionable health status of the test animals.

In some studies, especially with rabbits, there may be treatment with antimicrobial agents prior to use in teratology studies. Clearly, it is more desirable to use disease-free animals since the impact of drug treatment and disease on development are clearly unwanted variables. However, these studies may be considered acceptable on a case-by-case basis, depending upon the adequacy of the quarantine period prior to treatment and other factors.

Nulliparous animals should be selected for testing since confirmation of pregnancy in parous dams cannot be accurately determined. Furthermore, if corpora lutea are counted for the determination of preimplantation loss, nulliparous animals are preferred (WHO, Technical Report Series No. 364, 1967; See section on corpora lutea).

2. Dose-Levels

As stated in the Guidelines, at least three treatment levels and a control group (usually vehicle control) should be used. If there is any question as to toxicity of the vehicle, a sham or untreated control is also required. Although not required by the Guidelines, a concurrent positive control group may also be included. Dosing should be performed at the same time each day and reported.

Dosing at levels that produce excessive maternal toxicity may result in unreasonable numbers of deaths and abortions. Results such as these may limit the utility of the study and lead to study rejection. On the other hand, a study might also be rejected if the highest dose used is too low, i.e., without significant toxicity.

Although the Subdivision F Guidelines state that "the lowest dose level should not produce any evidence of toxicity," if maternal toxicity is the only finding at this dose-level, the study may still be considered acceptable.

3. Mortality and Clinical Observations

Understanding of the clinical signs of toxicity characteristic of the test compound should be gained from the range finding study and other toxicity studies prior to evaluation of the primary study.

Maternal death and/or abortion may be due to many factors, such as the test compound, diseases, environmental factors, and technical errors. Environmental factors, such as variations in housing conditions (temperature, humidity, light-cycle, caging, etc.), are known to influence the welfare of test animals. Also, technical errors such as intubation error and mishandling of animals in general can alter the outcome of test results and lead to maternal death and/or excessive stress.

The cause of death should be clarified from the necropsy data where possible. High incidences of congested lungs, reddening of the tracheal lining, and fluid accumulation in the lungs are suggestive of either technical error (e.g., gavage error) or disease. Consequently, care must be taken in the assessment of maternal death as an index of maternal toxicity since the two are often unrelated. The reviewer should always consider other possible reasons for maternal deaths (especially in rabbit studies). A dose-response relationship should be evident for one to conclude that effects are compound related to administration of the test compound.

4. Maternal Body Weight

Body weight and body weight gain data can be sensitive indicators of toxicity and are often used as a basis for the determination of the NOEL for maternal toxicity.

Test animals must be randomized for the body weight data to be useful as potential indicators of toxicity. This allows all dose groups to start with similar maternal weight means and variance.

Body weight measurements should be available at least on day 0 or 1 of gestation, at the start of dosing, on the final day of dosing, and on the day of sacrifice (usually one day prior to expected term).

Theoretically, the body weight gain (or percent change in body weight) of the treated groups should be comparable to that of the controls during the period prior to treatment. During the treatment period, a body weight reduction may be due to toxicity of the test compound and/or anorexia.

Anorexia is a common indicator of maternal toxicity and may occur soon after the initial dosing or may require repeated dosing before becoming evident. The effects of maternal anorexia may be assessed by calculating the food efficiency index which is

a measure of the efficacy of food utilization (grams of food consumed per kilograms of body weight gained).

The following table presents a typical pattern of compound related anorexia in a rabbit teratology study:

Assessment of a Typical Pattern of Food Intake During Gestation

Dose level mg/kg/day	Number of Pregnant Rabbits Available Per Interval*	Food Intake (g/dam/day) Gestation Days 0-5	6-18	19-27
0	14,13,10	106	76	53
250	13,13,10	115	56	79
500	15,14,4	102	23†	48
750	13,13,5	124	12†	26†

* Excluding spillers.
† Significantly different from control at $P < 0.05$.

Mean food intake for days 0-5 varied between 102 and 124 grams per day prior to administration of test compound. After test compound administration (days 6-18), a dose-related decrease in food consumption was observed in treated animals. The apparent decrease at the low dose-level was slight and within the normal range of variability. The mid and high dose-levels, however, were clearly less than the control values. After cessation of dosing (days 19-27), increased food consumption was observed in all treated groups as compared to the dosing period. This compensatory increase is common after a depression of food consumption during dosing. The overall pattern in food consumption indicates that compound-related effects were observed at the mid and high dose-levels during the period of dosing. The extent of the decrease clearly indicates maternal toxicity at these dose-levels. The low dose-level is comparable to the control group and, despite an apparent slight decrease, may be regarded as the No Observed Effect Level for maternal toxicity if there are no other indications of toxicity in the study (Burin, 1982).

If the food efficiency index is similar between treated and control groups, then anorexia may not be the main factor. Food consumption data are not presently required by the 1982 EPA Guidelines except for a dosed-feeding study. However, these data are considered useful for assessment of maternal effects regardless of route of administration.

Consequently, relevant interpretation of maternal toxicity may be obtained from the body weight gain data determined at different periods throughout gestation. Some of those periods are listed below:

a. Body weight gain throughout gestation
b. Body weight gain prior to dosing
c. Body weight gain during dosing
d. Body weight gain after dosing to study termination

Calculation of the "corrected" mean maternal weight gain (as measured by the difference in mean initial and terminal maternal body weight less the gravid uterus weight) may also serve as an index of maternal toxicity. This approach is considered more reliable by many investigators to assess maternal weight gain/loss, since it removes the uterine weight variability. An alternate but less desirable estimate of maternal weight change during gestation can be obtained by subtracting the litter weight from the maternal weight gain.

Assessment of maternal weight gain data in rabbits is difficult since erratic body weight gain is commonly found in this species. Nevertheless, comparison to concurrent control or historical control data may still provide the reviewer with valuable data for assessment.

Although reductions in maternal weight gain are generally dose-related, there are rare instances in which the low dose is affected while the high dose is not (Kavlock et al., 1984). This may be true when high maternal mortality is observed at high dose levels, eliminating sensitive animals, and consequently reducing the number of animals for comparison. In this case, the lack of a dose-response does not imply that there was no compound-related effect.

5. Organ Weights and Enzyme Markers

At the present time, neither organ weight data nor enzyme markers and clinical chemistry data are required by the Guidelines and as a result are often not available from studies for evaluation. However, dose-related effects on absolute and relative organ weights can be useful in the assessment of maternal toxicity. For example, the liver often shows the earliest signs of toxicity and therefore weight of this organ is occasionally reported and should receive careful consideration. Enzyme markers may sometimes provide additional data concerning exposure but must also be interpreted carefully as to whether or not a change constitutes toxicity.

Plasma, red blood cell and brain cholinesterase inhibition are rarely determined in a teratology study, but if these data are available and demonstrate pronounced effects (such as > 75% inhibition for red blood cell ChE), they may be used as a basis for establishing maternal toxicity (even in the absence of other clinical signs).

6. Conception Rate

The conception rate (# pregnant/# mated) is an important index for assessing the reproductive performance capability of the test animals selected. A low conception rate may suggest maternal health problems, poor animal husbandry, or may be due to dosing prior to completion of implantation. Regardless of the mechanism by which the conception rate is decreased, a reduction in available litters will result in a less meaningful interpretation of embryo/fetotoxicity (which includes assessment of teratogenic potential).

Both mating and artificial insemination have been used in rabbit teratology studies. Artificial insemination presumably will ensure a knowledge of the exact time of conception (Gibson et al., 1966). However, a high percentage of preimplantation loss and small litter size have been described with this method (Woo and Hoar, 1982; Woo, 1984). A conception rate of 80% and greater in rabbits is usually obtained by artificial insemination (Gibson et al., 1966, Adams, 1961; Woo, 1984) if the number and concentration of sperm used in the inseminating procedures are adequate (1 million sperm; Walton, 1927; Gibson et al., 1966). Artificial insemination techniques, semen dilution factors, and sperm motility are a few of the important parameters that should be included in the report. The time between luteinizing hormone and insemination is also critical since ova lose their fertilizability within 12 hours after hormonal treatment (Adams and Chang, 1962).

A conception rate of 88% or greater is usually obtained in rodents by pair or colony mating. Lower conception rates are common when pregnant animals are shipped from a supplier.

Since dosing in teratology studies is not to be initiated prior to the completion of implantation, as per the 1982 FIFRA Guidelines, no significant differences in conception rates between any test groups and controls should be observed. If significant differences are observed, the utility of the study may come into question and it may be rejected. Compound-related alterations in conception rate should generally be elucidated from data collected from multi-generation studies.

7. Corpora Lutea

The number of corpora lutea from both ovaries should be counted. Corpora lutea are evidence of released ova and generally outnumber implantation sites. Nulliparous animals are preferred in teratology studies to allow adequate determination of corpora lutea. This is to prevent potential counting error when distinguishing between corpora lutea and corpora albicans (from earlier ovulations).

Theoretically, the number of corpora lutea among test groups should be similar, since administration of test compound is to be performed after implantation has been completed.

8. Preimplantation Loss

Preimplantation loss is determined as the difference between the number of corpora lutea and implantation sites. An increase in preimplantation loss suggests a time-dosing error (dosing prior to the completion of implantation), which results in compound induced and/or maternal stress related embryolethality.

High preimplantation loss may obscure embryo/fetotoxicity and/or teratogenicity due to the reduced number of offspring left available for examination. In any case, an effort should be made to determine the cause of the high preimplantation loss.

MANIFESTATIONS OF DEVELOPMENTAL TOXICITY

The manifestations of developmental toxicity include death, structural abnormalities, altered growth, and functional disorder. Furthermore, there is an association between the manifestation observed and the developmental stage in which it is induced (Wilson and Fraser, 1977). Each type of manifestation is discussed in its respective section.

As previously indicated, developmental toxicity may or may not be related to maternal toxicity (Wilson and Fraser, 1977). For example, developmental toxicity and maternal toxicity are often observed at similar doses with alkylating agents, whereas thalidomide is devoid of maternal toxicity at doses that produce developmental toxicity. For organophosphate compounds, however, significant maternal toxicity is sometimes demonstrated at doses that result in little or no developmental toxicity. It should be obvious to the reviewer that agents of the greatest concern are those that induce developmental toxicity at levels significantly below those inducing maternal toxicity (also see "Review End Points to be Considered").

Developmental toxicity can be dependent upon the dosing schedule. Two dosing procedures can be utilized. The first limits dosing of the dams to the period of major organogenesis which is days 6 - 15 for rat and mouse, 6 - 14 for hamster, and 6 - 18 for rabbits. Alternatively, the dosing period may be extended to approximately 1 day before the expected day of delivery (Subdivision F Guidelines, 1982).

Exposure during the period of major organogenesis is based upon the principle that the embryo is most susceptible to induction of anomalies during this period; hence, exposure beyond this period is considered by some investigators to contribute little additional informative data. However, many organs such as the central nervous system, the heart, the lung, and the sexual organs, have not functionally or morphologically completed development by the end of major organogenesis. Consequently, other investigators believe that exposure of the conceptus from the day of implantation through gestation may disclose defects that normally would not be detected

under conventional teratology testing. It must be recognized that treatment of dams from the end of major organogenesis to sacrifice usually results in a higher incidence of fetotoxicity characterized mainly by altered growth and possibly fetal lethality. The reviewer should be aware that both dosing schedules are acceptable to the Office of Pesticide Programs.

1. Death

Depending on the stage of development affected, intrauterine death may be evident at sacrifice as resorptions and/or dead fetuses. These deaths may be related to maternal toxicity/stress or embryo/fetotoxicity including lethal malformations. These may be either spontaneous or compound-related. The total number of intrauterine deaths (resorptions and dead fetuses) is usually referred to as postimplantation loss or fetal wastage. Under certain circumstances where post-natal mortality data are available, these data should also be considered as part of this assessment.

Postimplantation loss is crucial to consider in assessing the embryo/fetotoxicity of an agent in these studies. In order to properly assess intrauterine death, fetuses should be recovered prior to expected delivery to prevent loss of prenatal material. Such losses frequently occur in many species after birth since the dams often cannibalize stillborn and/or abnormal pups. Furthermore, evidence of resorptions can only be accounted for from the maternal reproductive tract.

Before differentiation begins, the early embryo has great regenerative capability (totipotent), but if the dosage is sufficiently high, death of the conceptus may ensue. During early stages of organogenesis, susceptible embyros may die and be resorbed. This type of resorption is usually referred to as an early resorption which is generally complete or leaving only a very small amount of macerated tissue at the site of implantation or an implantation scar (metrial gland). As organogenesis progresses, toxic insults to the embryo may still result in resorption. The latter is referred to as a late resorption which is evidenced by the presence of both fetal tissue and placental tissue at the implantation site. After the period of major organogenesis, the fetus is increasingly resistant to the development of major malformations but still may display various types of fetotoxicity. Also, appreciable fetal lethality may still occur resulting in dead fetuses. Thus, evaluation of postimplantation loss (fetal wastage) takes into consideration both early and late resorption, as well as dead fetuses.

Analysis of postimplantation loss data may provide the most useful information regarding sensitive stages of development. This parameter is considered as a comprehensive expression of developmental toxicity since it can include all stages of development.

When embryolethal doses are reached, lethality increases at the expense of malformation. An increase in fetal wastage (post-implantation loss) may mask the teratogenic potential of an agent since affected malformed fetuses may be resorbed (Beck and Lloyd, 1963) and therefore unavailable for examination. Thus, in a teratology study, both a dose-related increase in fetal wastage (post-implantation loss) and malformations must be considered as indicative of embryo/fetotoxicity.

Only dead fetuses that do not show any significant degree of maceration should be included in the examination for malformations because of the frequent distortions and artifacts encountered in such specimens (Wilson, 1973).

One of the main problems in assessing fetal wastage is when pregnancy is suggested in a dam (proof of mating and subsequent weight gain) but no implantations and/or abortions are reported. In such case, it may not be possible to conclude whether or not embryo/fetotoxicity (including teratogenesis) has occurred or whether there has been an effect on fertility.

2. Altered Growth

In addition to known genetic, endocrine, and nutritional factors that influence growth, many agents which are capable of producing death and/or malformations may also cause altered growth. Two parameters usually employed to assess altered growth are fetal weight and fetal size (crown-rump length). The latter measurement is presently not required by the Guidelines.

It is generally assumed that fetal body weight is inversely proportional to litter size, i.e., higher fetal weight is usually expected from a smaller litter size and lower fetal weight is associated with larger litter size. Decreased fetal weight in conjunction with comparable litter size to controls is generally regarded as an indication of an embryo/fetotoxic effect.

Offspring that are two or three standard deviations below the mean control weight are classified as growth-retarded pups. If significant decreases in weight and length are found in one pup in comparison to the normal control range, this pup is then usually referred to as a runt. Runts are often considered manifestations of reversible embryo/fetotoxicity, but this is not always the case as is noted below.

Altered growth can be a transient or a permanent effect. In the case of a transient effect (reduced fetal weight, ossification retardation, etc.), a normal increase in weight, size, and maturation will be expected after birth. On the other hand, failure to recover from growth retardation is readily accepted as a defect (permanent stunting). The term runt is normally used to describe offspring exceptionally reduced in both size and weight. However, the criteria

used in making this distinction varies dramatically among laboratories (Palmer, 1977). Thus, the reversibility of growth retardation (including runting) must be assessed by considering appropriate data from conventional teratology, post-natal, and multi-generation reproduction studies. It should be emphasized that reproduction studies are required for all food-use pesticides and are useful for determining the reversibility of some aspects of embryo/fetotoxicity.

The reviewer should be aware that altered growth, like malformations, may only affect certain fetuses within a litter. Therefore if individual fetal data are available, altered growth should be assessed on both an individual and litter basis rather than solely on a mean litter basis.

Although effects on fetal sex ratio are quite rare, such data should always be examined since chemical agents may preferentially affect a particular sex (Scott et al., 1972). Other aspects of altered growth are discussed under the section on "anomalies."

3. Anomalies

The fetus is usually examined for external, soft tissue, and skeletal anomalies. Most investigators attempt to classify anomalies as either variations or malformations, although criteria for these classifications vary between laboratories. Variations are generally regarded as anomalies that may not adversely affect the fetuses and have no fatal outcome. Malformations, on the other hand, are anomalies that are considered to have a significant adverse effect on the fetus with or without fatal consequence. Distinction between variations and malformations is quite difficult in some cases.

Many skeletal anomalies are so common in laboratory animals that they are regarded as variations (if they do not have an adverse effect on the offspring). For example, changes in the numbers or degree of ossification of ribs and sternebrae in rabbits are considered variations (Gibson et al., 1966; Palmer, 1968; Woo, 1984). Although skeletal variations are not truly regarded as terata, they may be indicative of maternal toxicity/stress (Kavlock et al., 1984) and/or fetal toxic effects if a significant dose-related increase of a particular variation is noted above the concurrent controls. Historical control data should also be considered as part of this assessment.

Statistical analysis may support a determination as to whether results differ from those of the controls. In some aspects of assessment, N (on a litter basis) may be small and restricts the statistical sensitivity of the evaluation.

a. External Anomalies

All fetuses should be examined for external anomalies which are detected by means of a dissecting microscope. Limbs, tail,

axial skeleton, face, palate, eyes, and head are systematically examined.

An external malformation finding should be corroborated with soft tissue and/or skeletal examination data whenever possible.

Some examples of external malformations are listed below:

Club foot
Dome shaped head (Hydrocephaly)
Agnathia
Micrognathia
Spina bifida
Omphalocele
Gastroschisis
Phocomelia
Micromelia
Meningoencephalocele
Open eyes
Ectrodactyly
Syndactyly
Exencephaly
Cyclopia

As noted on page 2 under the definition of "malformations", no comprehensive list of malformations is possible.

b. Soft Tissue Anomalies

The Guidelines indicate that for rodents, approximately equal numbers of fetuses from each litter should be prepared and examined for skeletal anomalies and for soft tissue anomalies using appropriate methods. For rabbits, "each fetus should be examined by careful dissection for visceral anomalies and then examined for skeletal anomalies." The reviewer should understand that despite the Guidelines' reference to "visceral", the intention was that soft tissue examination data for both the head and torso are necessary. Deviations from the recommended protocol must be explained by the registrant.

The selected fetuses are examined by the method of Staples (1977) or fixed and examined for visceral anomalies by free-hand razor sectioning (Wilson, 1965). This sectioning technique has been found to be sufficiently adequate for rodent and rabbit fetuses.

Examples of some visceral anomalies are listed below:

Hydronephrosis
Dilated ureter
Anophthalmia
Diaphragmatic hernia
Ectocardia

Ectopic testis
Renal agenesis or ectopia

c. Skeletal Anomalies

The staining method most commonly used is the Dawson's technique (1926) using Alizarin Red S.

Distinctions between variations and malformations in skeletal development are sometimes problematic in animals most commonly used in teratology studies, such as the mouse, rat, hamster, and rabbit. As previously indicated, no universal distinction has been made between variations and malformations. Some investigators (see Palmer, 1968, 1972, 1977) divide skeletal anomalies into common variants, minor malformations, and major malformations on the basis of incidence and degree of deviation from "normal." Skeletal anomalies may be influenced by the genetic make-up of the species selected (Sawin et al., 1967), by maternal factors (Green, 1962), or compound induced (Palmer, 1972 and 1977).

Variations may be divided into groups such as delayed (retarded) ossification, extra ossification centers, or minor irregularities in shape and size.

As previous discussion has illustrated, skeletal anomalies should not necessarily be regarded as compound induced. However, the presence of a dose-related increase in the incidence of individual and/or anatomically-related skeletal variations is indicative of developmental toxicity (compound induced) and/or of a maternally related effect. For example, the presence of supernumerary ribs may be regarded as either embryo/fetotoxicity (Kimmel and Wilson, 1973) or maternally related effects (Kavlock et al., 1985).

Other examples of skeletal anomalies that are usually considered as skeletal variations include but are not restricted to, the following:

Skull, delayed ossification
Vertebrae, extra or absent ossification centers
Sternebrae, extra, absent, or incompletely ossified
Metacarpals (metatarsals), incomplete ossification or unossified
Ribs, extra, rudimentary, or incomplete ossified
Talus, incomplete ossification

The best rationale for interpretation of skeletal variations is comparison to concurrent and historical control data collected from various studies with the same species, strain of animals, and vehicle. Since there may be genetic drift in the strain of animals used, the reviewer must make a comparison with historical control data limited to a specific timeframe (that period being approximately 2 years prior to and possibly a comparable period subsequent to the

study). For each variation, both the number of pups and litters affected are important in making a final conclusion. Also, the pattern of variations should be considered. An increase in the incidence of variations can be assumed to be treatment related if the increase is statistically significant over the control range and particularly if the increase is dose-dependent. Trends should also be carefully considered.

Skeletal anomalies that are generally recognized as malformations include but are not restricted to:

Scoliosis
Kyphosis
Hemivertebrae
Syndactyly
Cranial vault, dome-shaped
Brachydactyly
Micromelia

As noted previously, the distinction between variations and malformations is sometimes quite difficult. In such cases, the reviewer should consult with a Toxicology Branch expert in this area. This distinction can be quite important since only an agent that causes a demonstrable increase in the incidence of malformations above the spontaneous rate (primarily concurrent but also historical control data) can be suspected of having a teratogenic effect.

Finally, it is critical for the reviewer to recognize two important facts (Palmer, 1977):

"1. That almost without exception every type of malformation ever recorded can occur sporadically in any species, and

2. That almost every type of malformation can arise from more than one cause."

In consideration of the above two points by Palmer, caution should be exercised by the reviewer in assessing teratogenic potential. For example, an individual occurrence of an extremely rare malformation in the absence of a dose-response curve may be spontaneous in origin. Conversely, if a larger population (N) were sampled in a repeat study, the possibility also exists that a treatment-related effect may be demonstrated.

4. Review End Points to be Considered

After following the assessment procedures presented in this document, the reviewer must make the following determinations:

Maternal Toxicity

As stated previously in this SEP, maternal toxicity must be demonstrated at the high dose-level and a No Observed Effect Level

(NOEL) and Lowest Observed Effect Level (LOEL) should be determined by the reviewer. If the reviewer cannot determine a NOEL for maternal toxicity even at the lowest dose, as long as a NOEL for developmental toxicity (embryo/fetotoxicity) is available from this study, a repeat study is not necessary.

° Developmental Toxicity

As defined previously, developmental toxicity includes death of the developing organism, structural abnormalities (malformations and variations), altered growth, and functional deficiency. At present, the latter end point is not assessed in a conventional Guidelines teratology study but useful data is sometimes available from reproductive studies and post-natal studies (if required). These additional studies (and especially post-natal studies) also offer useful information for assessing the reversibility of altered growth as well as providing data on other aspects of developmental toxicity. It must also be understood that in a classical teratology study, developmental toxicity is restricted to adverse effects manifested in the developing organism prior to parturition. Therefore, NOELs for adverse effects which were observed in classical teratology studies (studies which encompass the embryonic and fetal periods of the organism) are more accurately referred to as "developmental toxicity (embryo/fetotoxicity) NOELs" while adverse effects which are observed in post-natal studies should be referred to as "developmental toxicity (post-natal) NOELs."

a. Embryo/Fetotoxicity

Assessment of developmental toxicity (embryo/fetotoxicity) must include determinations as to whether there are dose-related increases in postimplantation loss [early resorptions (embryolethality), plus late resorptions, plus death] as well as variations, fetal body weight changes and fetal crown-rump length changes (when available). As stated previously in the definitions section of this SEP (page 2), the reviewer should also be aware that developmental toxicity (embryo/fetotoxicity) includes malformations. Therefore, the NOEL and LOEL determined by the reviewer should be based on all of the above end points and a separate NOEL for teratogenicity (as has been customary in the past) will not be necessary. Malformation and fetal wastage data will now be presented in text form as described below. If no developmental toxicity (embryo/fetotoxicity) NOEL can be determined, a repeat study may be required.

b. Teratogenic Potential

The reviewer must determine whether a study is positive or negative for teratogenicity but this will not necessitate a separate teratogenic NOEL and LOEL. As discussed earlier, since teratogenicity is only a subset of developmental toxicity, distinction between NOELs for subsets of developmental toxicity may not be relevant.

The concept of developmental toxicity is not a new one. Wilson (1973) indicated that there are four signs of developmental toxicity: 1) death and resorption, 2) birth of either live or dead malformed offspring, 3) developmental delays, and 4) decrement of expected post-natal function. All four end points are of toxicological concern if produced at a particular exposure level and in a dose-related manner (Johnson and Christian, 1984). For example, in utero death or resorptions are as significant as production of live but malformed offspring. Therefore, the reviewer should indicate:

- A developmental toxicity (embryo/fetotoxicity) NOEL and and LOEL for all teratology studies following § 83-3 of the Guidelines.

- If a dose-related increase in the incidence of malformations was noted and/or fetal wastage (postimplantation loss) was observed, the reviewer, in addition to the developmental toxicity (embryo/fetotoxicity) NOEL and LOEL, should accurately describe in the conclusions/recommendations section of the review the types of effects and the dose-levels at which these adverse effects occurred.

° Developmental Toxicity Index

A number of investigators have suggested that a useful way to assess the relative developmental hazard is through calculation of a developmental toxicity index.

Johnson (1980; 1981) describes agents as "coeffective teratogens" if they affect the embryos only at doses near the adult toxic range and as "non-coeffective" if effects are observed in the offspring at dose-levels below the adult toxic range. Using the above terminology, it is the "non-coeffective" developmental toxins which are usually, but not always, of greatest concern to the Agency.

Quantitative estimates of developmental hazards have been proposed by several investigators. The "Relative Teratogenic Index" was introduced by Fabro et al. (1982, 1985). These investigators proposed a quantitative estimate of teratogenic potency using the "Relative Teratogenic Index" which is a calculated ratio of the maternal adult toxic dose (LD_{01}) and the tD_{05} (minimum teratogenic dose based upon dose-response analysis of teratogenicity fitted to a probit model). Johnson and Gabel (1982) suggested the use of the A/D ratio which is the minimal effective concentration toxic to the adult (A) and to the developing organism (D). However, it should be noted that although all proposed estimates of relative developmental hazard are of merit, none of them should automatically be applied and used by the reviewer without proper consideration of the limitations of the data. It should be recognized that the utility of all of these index systems is limited by (1) the dose selection (both spacing of the dose-levels and the few levels actually tested) which is used to define maternal and developmental toxicity, (2) the limited number of parameters used for assessing maternal and developmental toxicity, (3) species selection and number of animals tested

per group, (4) route of administration, etc.. It must also be understood that these calculations are indices of hazard but do not represent measures of risk since they do not consider levels of human exposure. In conclusion, it is important for the reviewer to be aware of the many limitations of the A/D ratio and other index systems.

The A/D ratio (Johnson and Gabel, 1982) takes into consideration the maternal LOEL and developmental toxicity (embryo/fetotoxicity) LOEL. A ratio of less than 1 indicates that developmental toxicity occurs at doses higher than those producing maternal toxicity. A ratio of 1 indicates that developmental effects are found at doses which also produced maternal toxicity. However, a ratio of more than 1 reveals that developmental toxicity occurs at doses lower than those producing maternal toxicity. Using this approach, the higher A/D ratio would suggest a higher developmental hazard for that species, route, etc..

° Acceptability of the Study for Regulatory Purposes

The guidance provided in this document is intended to be used in conjunction with the Core Classification System for determining the acceptability of a teratology study for regulatory purposes.

5. Functional Defects - Assessment of Post-Natal Studies

Developmental toxicity includes any developmental changes induced at any stage of gestation and detected not only at birth but also at any time post-natally. The protocol, as listed in the Subdivision F Guidelines, is restricted to an examination at cesarean section. Consequently, functional changes in the offspring (and even viability) cannot be fully assessed from the Guidelines protocol. Furthermore, there are data which demonstrate that some teratogenic manifestations of developmental toxicity which appear after birth are not detected by the classical Guidelines study (Gray et al., 1982; Chernoff and Kavlock, 1982; Bui et al., 1983).

If assessment of the classical study data raises any question relative to the permanent status of an abnormality or raises other concerns, the reviewer should consider requesting post-natal studies. Other supportable rationales may also be used as a basis for requesting these studies. Examples of dose-related findings that may trigger such requests include (but are not restricted to): dilated renal pelvis and/or dilated ureters, runting, and abnormalities of the central nervous system.

A reviewer should consider data from the reproduction study in assessment of possible post-natal effects; however, one must recognize that (1) the route of exposure is usually via the diet rather than gavage, (2) the dose-levels are usually lower and extended over longer periods of time, and (3) these studies do not include the level of fetal examination employed in classical teratology studies. Thus,

although these data should be considered, they can rarely provide an adequate assessment of all potential post-natal effects.

As stated previously, the NOEL and LOEL for the adverse effects which are observed in a post-natal study should be referred to as "developmental toxicity (post-natal)" NOELs and LOELs.

In the design of acceptable protocols for post-natal studies, the reviewer should consult with Toxicology Branch experts. Protocols will vary widely and studies may in some cases be terminated at weaning, at sexual maturation, or even later.

IV. DEVELOPMENTAL TOXICITY RISK ASSESSMENT

Developmental toxicity risk assessment includes consideration of all relevant data such as pharmacokinetic, metabolism, structure-activity relationships (although not well studied) and other studies.

Many methods have been proposed for the extrapolation of developmental toxicity risk from test animals to humans and this area continues to be an active one for research and scientific discussion (Hogan, 1982). The "Margin of Safety" and the closely related "Safety Factor" approaches are widely used and accepted. The Office of Pesticide Programs has used the Margin of Safety approach to evaluate potential developmental toxicity risk (e.g., Chitlik, 1980). It has recently been cited in the Agency's "Interim Guidelines for the Health Assessment of Suspect Developmental Toxicants" as a recommended approach that the Agency use for developmental toxicity risk assessment. The primary difference between the two approaches is the *a priori* consideration of exposure data in the Margin of Safety approach. Thus, the Margin of Safety approach is a direct comparison (ratio) between the appropriate No Observed Effect Level (NOEL) and the estimated human exposure, while the Safety Factor approach divides the NOEL by a somewhat arbitrary numerical value (the Safety Factor) to reach a "safe" level of exposure.

If a concern for the effects of developmental toxicity as indicated by this SEP is suggested by one or more scientifically sound studies, the reviewer should determine the No Observed Effect Level in the most sensitive species tested. The most sensitive species is generally used for developmental toxicity risk assessment purposes due to the great difficulty in determining the most relevant species from which to extrapolate to humans (USEPA, 1984a).

Ideally, epidemiological studies of birth defects after exposure to pesticides would be the most relevant type of study upon which to estimate risk. These studies, if available, are usually inadequate for quantitative risk assessment due to many inherent limitations which are outside the scope of this document.

The reviewer should next examine the types and levels of actual or potential human exposure, keeping in mind that develop-

mental toxicity may result from a single exposure. Developmental toxicity risk assessment routinely includes dietary and worker exposure, as well as other forms of exposure such as drinking water or home use. Particular attention should be given to worker exposure as it is frequently a much greater acute exposure than that received via the dietary route. Worker exposure estimates are the responsibility of the Exposure Assessment Branch (EAB). If a worker exposure estimate is not available, the reviewer should consult with the Section Head for guidance on requesting an exposure assessment to be conducted by EAB. The EAB estimates of exposure must be on a daily basis and should be quantified for each route of exposure.

It is the responsibility of Toxicology Branch to determine the rate of dermal absorption. In the proper assessment of potential risk from dermal or other routes of exposure, one should also compare pharmacokinetic data such as peak plasma concentrations and/or area under the curve of the test material and/or metabolites when dosing is via different routes. This may prove important in understanding differences in developmental toxicity potential to workers (primarily dermal and inhalation versus dietary exposure hazards; Chitlik and Bui, 1985). Metabolism data should also be considered since this may vary somewhat with route of exposure. In the absence of dermal absorption data, a 100% rate of absorption should be assumed.

Dietary exposure must also be assessed on a daily basis. The standard "food factor" approach (Schmitt, 1978) does not provide information regarding expected daily exposure but is an annualized average. When completed, the new Tolerance Assessment System (USEPA, 1984b) will not have this limitation and will be the primary source for reviewers to obtain this information. Another useful source for estimating single serving size is the U.S. Department of Agriculture's "Family Food Buying Guide" (USDA, 1977). Other references or rationales may also be used in the determination of single serving size. Regardless of the source used, the reviewer must estimate the maximum amount (kg) of a raw agricultural commodity that is likely to be consumed on a given day by an individual.

The USDA guide provides the reviewer with the "size of market unit" as well as the "number of servings or measures per market unit." The reviewer can then calculate from this information a number of conceivable single serving sizes and should use, for risk assessment purposes, the largest value. A total dietary exposure can then be estimated by multiplying the residue level (expressed in ppm) by the maximum likely serving size per day (in kg) and dividing this by the weight of the individual exposed. The Toxicology Branch has historically used 60 kg as the body weight of females of child bearing age for developmental toxicity assessment. As per the new Tolerance Assessment System (TAS), females of child bearing age are estimated to weigh 54.8 kg. The appropriate residue level is that which is recommended by the Residue Chemistry Branch and may either be at the tolerance level or at the actual exposure

level (if such data are made available).

The reviewer must utilize judgment in selecting the relevant dietary exposure estimates. The reviewer should consider individual raw agricultural commodities as well as realistic combinations of food items which may be consumed in a single day. In all cases, the reviewer should attempt to be conservative but not unrealistic in the exposure assumptions made in the risk assessment.

Drinking water risks may also be of concern and are assessed in a manner similar to dietary risk. Determination of whether or not a pesticide has a potential for groundwater or surface water contamination is the responsibility of the Exposure Assessment Branch. If actual contamination exists, the most relevant contamination levels must be selected in consultation with the EAB. The National Academy of Sciences has recommended that, for risk assessment purposes, it be assumed that the average adult consumes two liters of water per day (Office of Drinking Water/EPA and NAS, 1977). The estimated daily exposure (mg/kg/day) to a pesticide in drinking water is therefore determined by multiplying the appropriate estimate of the residue level (mg/liter) by two liters and dividing that amount by body weight. A more detailed description of the hazard evaluation of pesticides in drinking water can be found in the 1983 document "Assessment of Groundwater Contamination by Pesticides" (USEPA, 1983) and in the Federal Register of June 12, 1984.

The exposure estimates generated for the worker and the consumer of food or water containing the chemical of concern can then be compared as a ratio to the No Observed Effect Level for the relevant effect to yield an estimate of the Margin of Safety (MOS). When possible, pharmacokinetic and metabolic considerations should be taken into account in extrapolation from one route of administration to another.

This determination is useful to the risk managers who must balance the health risks against societal costs and benefits in the determination of appropriate regulatory action. The Margin of Safety approach is a means of relating the potency of an agent to induce an effect with an estimate of human exposure.

In a modified form, the MOS is useful even when a NOEL cannot be assessed from available data. In that case a "what if" scenario can illustrate possible Margins of Safety associated with speculated (or hypothetical) NOELs (Chitlik, 1983). The following table illustrates the effect of a reduction in a potential NOEL from 0.1 to 0.05 to 0.01 mg/kg/day upon the calculated Margin of Safety:

Food Item	Serving Size (kg)[a]	Tolerance (ppm)	Exposure (mg/kg bw/ day) [b]	MOS assuming developmental tox. NOEL of : 0.1, 0.05, 0.01 (mg/kg/day)		
Broccoli	.103	0.10	0.00018	543	271	54
Brussel-Sprouts	.081	0.75	0.00109	92	46	9
Cabbage	.140	0.75	0.00188	34	17	3
Celery	.170	0.10	0.00030	328	164	33
Cauliflower	.140	0.30	0.00075	133	67	13
Kohlrabi	.182	0.75	0.00244	41	20	4
Onions (Dry bulbs)	.202	0.75	0.00271	37	18	4
Sugar beets (sugar)	.202	0.05	0.00018	554	277	55

(a) Size of individual food services were obtained from Family Food Buying, L. Fultons C. Davis and E. Matthews, USDA No.37, 08/1977.

(b) female body weight of 55.9 kg utilized based upon TAS draft

(Note: now revised to 54.8 kg; USEPA, 1984b)

Calculations of Margin of Safety must be qualified by description of the strength of the evidence and biological uncertainties. The National Research Council, in a widely cited report "Risk Assessment in the Federal Government: Managing the Process" has recommended that risk assessments "... set forth in detail the nature and quality of the relevant scientific evidence concerning the substance in question and should cover all relevant components of risk assessment" (NRC, 1983). For developmental risk assessment, the uncertainties are great and have been discussed in detail in several Agency documents, e.g., USEPA, 1980, and USEPA, 1984a. The routine assumptions made in the extrapolation of animal data to humans need not be reiterated for each risk assessment. The reviewer should concentrate on the discussion of the quality of the studies supporting the concern for developmental toxicity risk, the strength of the NOEL and the uncertainties in the exposure assessment (such as the lack of information regarding dermal absorption or uncertainties relative to residue levels in the diet).

V. BIBLIOGRAPHY

Adams, C.E., 1961. Artificial insemination. In: The semen of animals and artificial insemination. ed. J.P. Maude. England.

Benirschke, K., Garner, F.M., and Jones, T.C., 1978. In: Pathology of Laboratory Animals, eds. K. Benirschke, F.M. Garner, and T.C. Jones, Springer Verlag.

Beck, F. and Lloyd, J.B., 1963. An investigation of the relationship between foetal death and foetal malformation. J. Anat. 97, pp. 555-564.

Bui, Q.Q., Sperling, F. and West, W.L., 1983. Developmental toxic effect after subcutaneous injections of methadone in Charles River CD-1 mice. Drug and Chemical Toxicology, 6(1), pp. 41-70.

*Burin, G.J., 1982. Review of rabbit teratology studies of Ethalfluralin. Memorandum to William Burnam, 11/24, USEPA.

Chang, M.C., 1958. Capacitation of rabbit spermatozoa in the uterus with special referene to the reproductive phases of the female. J. Endocrinol. 63, pp. 619-628.

Chernoff, N. and Kavlock, R.J., 1982. An in vivo teratology screen utilizing pregnant mice. J. Toxicol. Environm. Health 10, pp. 541-550.

*Chitlik, L.D., 1980. TOK (Nitrofen) - Oncogenic, teratogenic, and mutagenic evaluations. Memorandum to Marcia Williams, 05/21, USEPA.

*Chitlik, L.D., 1983. Margin of Safety (MOS) assessment for Nitrofen (TOK). Memorandum to M. Branagan, 06/15, USEPA.

*Chitlik, L.D. and Bui, Q.Q., 1985. Review of toxicology studies with Karathane. Memorandum to Jay Ellenberger, 04/01, USEPA.

Dawson, A.B., 1926. A note on the staining of the skeleton of cleared specimens with Alizarin Red S, Stain Technology 1, p. 123.

Fabro, S., 1985. On predicting environmentally-induced human reproductive hazards: An overview and historical perspective. Fund. Appl. Toxicol. 5, pp. 609-614.

Fabro, S., Shull, G. and Brown, N.A., 1982. The relative teratogenic index and teratogenic potency: Proposed components of developmental toxicity risk assessment. Terato. Carcino. and Mutagenesis 2, pp. 61-76.

Federal Register, 1984. Volume 49, No. 114, pp. 24330-24335.

Gibson, J.P., Staples, R.E., and Newberne, J.W., 1966. Use of rabbit in teratogenicity studies. Toxicol. Appl. Pharmacology 9, pp. 398-408.

Gray, L.E. Jr., Kavlock, R.J., Chernoff, N., et al., 1982. Prenatal exposure to the herbicide 2,4-dichlorophenyl-p-nitrophenyl ester destroys the rodent harderian gland. Science 215, pp. 293-294.

Green, E.L., 1962. Quantitative genetics of skeletal variations in the mouse. II. Crosses between four inbred strains (C3H, DBA, C57BL, BALB/c). Genetics 47: pp. 1085-1096.

Hafez, E.S.E., 1970. Reproduction and breeding techniques for laboratory animals, eds. Lea and Febiger, Philadelphia.

Hogan, M.D. and Hoel, D.G., 1982. Extrapolation to man. In: Principles and methods of toxicology. ed., A.W. Hayes, Raven Press.

Johnson, E.M., 1980. A subvertebrate system for rapid determination of potential teratogenic hazards. J. Environ. Pathol. Toxicol. 4, pp. 153-156.

Johnson, E.M., 1981. Screening for teratogenic hazards: Nature of the problems. Annu. Rev. Pharmacol. Toxicol. 21, 417-429.

Johnson, E.M. and Gabel, B.E.G., 1982. Application of the hydra assay for rapid detection of developmental hazards. J. Amer. Col. Toxicol. 1(3), pp. 57-71.

Johnson, E.M. and Christian, M.S., 1984. When is a teratology study not an evaluation of teratogenicity? J. Amer. Col. Tox. 3, pp. 431-434.

Kavlock, R.J., Chernoff, N., and Rogers, E.H., 1984. Supernumerary ribs: Is there a relationship with alterations in maternal health status. Teratogenesis Carcino. and Mutagen. 5(1).

Kavlock, R.J., Chernoff, N., and Rogers, E.H., 1985. The effect of acute maternal toxicity on fetal development in the mouse. Teratogenesis, Carcino. and Mutagen. 5, pp. 3-13.

Kimmel, C.A. and Wilson, J.G., 1973. Skeletal deviations in rats: Malformations or variations? Teratology 8, pp. 309-316.

National Academy of Sciences, July 11, 1977. Drinking water and health, Federal Register, 42, p. 132.

National Research Council, 1983. Risk assessment in the federal government: Managing the process. National Academy Press, Washington D.C., p. 159.

Palmer, A.K., 1968. Spontaneous malformations of the New Zealand white rabbit: The background to safety evaluation studies. Lab. Animals 2, pp. 195-206.

Palmer, A.K., 1972. Sporadic malformations in laboratory animals and their influence on drug testing. In: Drugs and Fetal Development eds M.A. Klinberg, A. Abramovici, and J. Chemke. Plenum Press.

Palmer, A.K., 1977. Incidence of sporadic malformations, anomalies and variations in random bred laboratory animals. In: Methods in Prenatal Toxicology. eds D. Neubert, H.J. Merker, and T.E. Kwasigroch Georg Thieme Pub., Stuggart.

Sawin, P.B. and Gow, M. 1967. Morphogenetic studies of the rabbit. XXXVI. Effect of gene and genome interaction on homeotic variation Anat. Rec. 157, pp. 425-435.

Schardein, J.L., 1985. In: Chemically induced birth defects. ed. J.L. Schardein, Marcel Dekker Pub., New York

Schmitt, D., 1978. Update of food factor tables. Toxicology Branch, HED, OPP, USEPA.

Scott, W.J., Butcher, R.L., Kindt, C.W., and Wilson, J.G., 1972 Greater sensitivity of female than male rat embryos to acetazolamide teratogenicity. Teratology 6, pp. 239-240.

Shepard, T.H., 1983. Catalog of Teratogenic Agents. John Hopkins University Press.

Staples, R.E., 1977. Detection of visceral alterations in mammalian fetuses. Teratology 9, pp. A37 - A38.

U.S. Dept. of Agriculture, 1977. Family food buying guide.

U.S. Dept. of Agriculture, 1982. Foods commonly eaten by individuals; Amount for day and per eating. Human Nutritional Information Service. Home Econ. Res. Rep. 44.

USEPA, 1980. Assessment of risks to human reproduction and to development of the human conceptus from exposure to environmental substances. National Technical Information Service US. Dept. of Commerce, No. DE 82-007897.

USEPA, 1982. Pesticide Assessment Guidelines, Subdivision F, Hazard Evaluation: Human and Domestic Animals.

USEPA, 1983. Assessment of ground water contamination by pesticides Hazard Evaluation Division, OPP, USEPA.

USEPA, 1984a., Proposed Guidelines for the health assessment of suspect developmental toxicants and request for comments. Federal Register, November 23.

USEPA, 1984b., Ethylene Dibromide (EDB) - Scientific support and Decision. Document for grain and grain milling fumigation uses. Office of Pesticide Programs. pp. 47-49.

Walton, A., 1927. Relation between density of sperm suspension and fertility as determined by artificial insemination of rabbits. Proc. Roy. Soc. B 101, pp. 303-315.

WHO, 1967. Principles for the testing of drugs for teratogenicity. WHO Technical Report Series No. 364, Geneva, Switzerland.

Wilson, J.G., 1965. Methods for administering agents and detecting malformations in experimental animals. In: Teratology: Principles and Techniques., eds J.G. Wilson and J. Warkany. Univ. of Chicago Press.

Wilson, J.G. 1973. In: Environment and birth defects. ed., J.G. Wilson, Academic Press.

Wilson J.G. and Fraser, F.C., 1977. in :Handbook of teratology. I. General Principles and Etiology. eds. J.G. Wilson and F.C. Fraser, Plenum Press.

Woo, D.C. 1984. Control data of teratological studies in New Zealand White rabbits (A joint study by MARTA), in Principles in Teratology, Teratology Society.

Woo, D.C. and Hoar, R.M., 1972. "Apparent hydronephrosis" as a normal aspect of renal development in late gestation of rats: The effect of methyl salicylate. Teratology 6, pp. 191-196.

Woo, D.C. and Hoar, R.M., 1982. Reproductive performance and spontaneous malformations in control New Zealand rabbits: A joint study by MARTA. Teratology 25, p. 82A.

* These references are internal EPA documents which contain confidential data. Removal of such data will eliminate the instructional value of the document. Therefore, these documents must not be released from HED without the proper authorization and legal clearances.